

ABSTRACT

A system and method for improving the accuracy of DNA sequencing and error probability estimation through application of a mathematical model to the analysis of electropherograms. The method includes processing a plurality of information obtained from a base calling system and creating a plurality of refined base calls using a plurality of original base calls and a plurality of intrinsic peak characteristics. A quality value is also assigned to each of the plurality of refined base calls using the plurality of intrinsic peak characteristics. Processing comprises detecting a plurality of peaks, expanding the plurality of peaks, and resolving the plurality of expanded peaks. Resolving may include fitting the plurality of expanded peaks using a model of a peak shape. A peak resolution parameter is calculated and used in processing. The system may also be trained.